# OIPE

*INPUT SET: S33766.raw*

This Raw Listing contains the General
Information Section and up to the first 5 pages.

ENTERED

SEQUENCE LISTING

```
 1
 2
 3   (1)     General Information:
 4
 5    (i) APPLICANT: Das, Goutam
 6
 7   (ii) TITLE OF INVENTION: DNA Molecules for Expression of
 8            Polypeptides
 9
10  (iii) NUMBER OF SEQUENCES: 4
11
12   (iv) CORRESPONDENCE ADDRESS:
13            (A) ADDRESSEE: White & Case
14            (B) STREET: 1155 Avenue of the Americas
15            (C) CITY: New York
16            (D) STATE: New York
17            (E) COUNTRY: United States
18            (F) ZIP: 10036-2787
19
20    (v) COMPUTER READABLE FORM:
21            (A) MEDIUM TYPE: Floppy disk
22            (B) COMPUTER: IBM PC compatible
23            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
24            (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
25
26   (vi) CURRENT APPLICATION DATA:
27            (A) APPLICATION NUMBER:
28            (B) FILING DATE:
29            (C) CLASSIFICATION:
30
31  (vii) PRIOR APPLICATION DATA:
32            (A) APPLICATION NUMBER: 08/624,398
33            (B) FILING DATE: 04-APR-1996
34
35  (vii) PRIOR APPLICATION DATA:
36            (A) APPLICATION NUMBER: PCT/SE96/00318
37            (B) FILING DATE: 12-MAR-1996
38
39  (vii) PRIOR APPLICATION DATA:
40            (A) APPLICATION NUMBER: SE 9501939-4
41            (B) FILING DATE: 24-MAY-1995
42
43 (viii) ATTORNEY/AGENT INFORMATION:
44            (A) NAME: Thelma A. Chen Cleland
45            (B) REGISTRATION NUMBER: 40,948
46            (C) REFERENCE/DOCKET NUMBER: 1103326-0206
```

RECEIVED NOV -3 1999 1700 MAIL ROOM

```
47
48          (ix)  TELECOMMUNICATION INFORMATION:
49                (A)  TELEPHONE:  (212) 819-8200
50                (B)  TELEFAX:  (212) 354-8113
51
52
53    (2)  INFORMATION FOR SEQ ID NO:1:
54
55          (i)  SEQUENCE CHARACTERISTICS:
56                (A)  LENGTH: 2428 base pairs
57                (B)  TYPE: nucleic acid
58                (C)  STRANDEDNESS: double
59                (D)  TOPOLOGY: linear
60
61         (ii)  MOLECULE TYPE: cDNA to mRNA
62
63        (iii)  HYPOTHETICAL: NO
64
65         (iv)  ANTI-SENSE: NO
66
67         (vi)  ORIGINAL SOURCE:
68                (A)  ORGANISM: Homo sapiens
69                (F)  TISSUE TYPE: mammary gland
70
71         (ix)  FEATURE:
72                (A)  NAME/KEY: CDS
73                (B)  LOCATION: 82..2319
74                (D)  OTHER INFORMATION: /product= "bile-salt-stimulated
75    lipase"
76
77         (ix)  FEATURE:
78                (A)  NAME/KEY: exon
79                (B)  LOCATION: 985..1173
80
81         (ix)  FEATURE:
82                (A)  NAME/KEY: exon
83                (B)  LOCATION: 1174..1377
84
85         (ix)  FEATURE:
86                (A)  NAME/KEY: exon
87                (B)  LOCATION: 1378..1575
88
89         (ix)  FEATURE:
90                (A)  NAME/KEY: exon
91                (B)  LOCATION: 1576..2415
92
93         (ix)  FEATURE:
94                (A)  NAME/KEY: mat_peptide
95                (B)  LOCATION: 151..2316
96
97         (ix)  FEATURE:
98                (A)  NAME/KEY: polyA_signal
99                (B)  LOCATION: 2397..2402
```

PAGE: 3

**RAW SEQUENCE LISTING**
**PATENT APPLICATION**   *US/09/418,176*

DATE: 10/26/1999
TIME: 13:34:28

*INPUT SET: S33766.raw*

```
100
101        (ix)  FEATURE:
102              (A)  NAME/KEY:  repeat_region
103              (B)  LOCATION:  1756..2283
104
105        (ix)  FEATURE:
106              (A)  NAME/KEY:  5'UTR
107              (B)  LOCATION:  1..81
108
109        (ix)  FEATURE:
110              (A)  NAME/KEY:  repeat_unit
111              (B)  LOCATION:  1756..1788
112
113        (ix)  FEATURE:
114              (A)  NAME/KEY:  repeat_unit
115              (B)  LOCATION:  1789..1821
116
117        (ix)  FEATURE:
118              (A)  NAME/KEY:  repeat_unit
119              (B)  LOCATION:  1822..1854
120
121        (ix)  FEATURE:
122              (A)  NAME/KEY:  repeat_unit
123              (B)  LOCATION:  1855..1887
124
125        (ix)  FEATURE:
126              (A)  NAME/KEY:  repeat_unit
127              (B)  LOCATION:  1888..1920
128
129        (ix)  FEATURE:
130              (A)  NAME/KEY:  repeat_unit
131              (B)  LOCATION:  1921..1953
132
133        (ix)  FEATURE:
134              (A)  NAME/KEY:  repeat_unit
135              (B)  LOCATION:  1954..1986
136
137        (ix)  FEATURE:
138              (A)  NAME/KEY:  repeat_unit
139              (B)  LOCATION:  1987..2019
140
141        (ix)  FEATURE:
142              (A)  NAME/KEY:  repeat_unit
143              (B)  LOCATION:  2020..2052
144
145        (ix)  FEATURE:
146              (A)  NAME/KEY:  repeat_unit
147              (B)  LOCATION:  2053..2085
148
149        (ix)  FEATURE:
150              (A)  NAME/KEY:  repeat_unit
151              (B)  LOCATION:  2086..2118
152
```

*INPUT SET: S33766.raw*

```
153        (ix)  FEATURE:
154              (A)  NAME/KEY: repeat_unit
155              (B)  LOCATION: 2119..2151
156
157        (ix)  FEATURE:
158              (A)  NAME/KEY: repeat_unit
159              (B)  LOCATION: 2152..2184
160
161        (ix)  FEATURE:
162              (A)  NAME/KEY: repeat_unit
163              (B)  LOCATION: 2185..2217
164
165        (ix)  FEATURE:
166              (A)  NAME/KEY: repeat_unit
167              (B)  LOCATION: 2218..2250
168
169        (ix)  FEATURE:
170              (A)  NAME/KEY: repeat_unit
171              (B)  LOCATION: 2251..2283
172
173        (x)  PUBLICATION INFORMATION:
174             (A)  AUTHORS: Nilsson, Jeanette
175                           Blackberg, Lars
176                           Carlsson, Peter
177                           Enerback, Sven
178                           Hernell, Olle
179                           Bjursell, Gunnar
180             (B)  TITLE: cDNA cloning of human-milk
181                  bile-salt-stimulated lipase and evidence for its
182                  identity to pancreatic carboxylic ester hydrolase
183             (C)  JOURNAL: Eur. J. Biochem.
184             (D)  VOLUME: 192
185             (F)  PAGES: 543-550
186             (G)  DATE: Sept.-1990
187
188        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:1:
189
190
191   ACCTTCTGTA TCAGTTAAGT GTCAAGATGG AAGGAACAGC AGTCTCAAGA TAATGCAAAG         60
192
193   AGTTTATTCA TCCAGAGGCT G ATG CTC ACC ATG GGG CGC CTG CAA CTG GTT        111
194                         Met Leu Thr Met Gly Arg Leu Gln Leu Val
195                         -23           -20                 -15
196
197   GTG TTG GGC CTC ACC TGC TGC TGG GCA GTG GCG AGT GCC GCG AAG CTG        159
198   Val Leu Gly Leu Thr Cys Cys Trp Ala Val Ala Ser Ala Ala Lys Leu
199               -10               -5                    1
200
201   GGC GCC GTG TAC ACA GAA GGT GGG TTC GTG GAA GGC GTC AAT AAG AAG        207
202   Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly Val Asn Lys Lys
203        5                 10                 15
204
205   CTC GGC CTC CTG GGT GAC TCT GTG GAC ATC TTC AAG GGC ATC CCC TTC        255
```

*INPUT SET: S33766.raw*

```
206    Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly Ile Pro Phe
207     20              25              30              35
208
209    GCA GCT CCC ACC AAG GCC CTG GAA AAT CCT CAG CCA CAT CCT GGC TGG    303
210    Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His Pro Gly Trp
211                 40              45              50
212
213    CAA GGG ACC CTG AAG GCC AAG AAC TTC AAG AAG AGA TGC CTG CAG GCC    351
214    Gln Gly Thr Leu Lys Ala Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala
215                 55              60              65
216
217    ACC ATC ACC CAG GAC AGC ACC TAC GGG GAT GAA GAC TGC CTG TAC CTC    399
218    Thr Ile Thr Gln Asp Ser Thr Tyr Gly Asp Glu Asp Cys Leu Tyr Leu
219                 70              75              80
220
221    AAC ATT TGG GTG CCC CAG GGC AGG AAG CAA GTC TCC CGG GAC CTG CCC    447
222    Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser Arg Asp Leu Pro
223           85              90              95
224
225    GTT ATG ATC TGG ATC TAT GGA GGC GCC TTC CTC ATG GGG TCC GGC CAT    495
226    Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly Ser Gly His
227    100             105             110             115
228
229    GGG GCC AAC TTC CTC AAC AAC TAC CTG TAT GAC GGC GAG GAG ATC GCC    543
230    Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala
231                 120             125             130
232
233    ACA CGC GGA AAC GTC ATC GTG GTC ACC TTC AAC TAC CGT GTC GGC CCC    591
234    Thr Arg Gly Asn Val Ile Val Val Thr Phe Asn Tyr Arg Val Gly Pro
235                 135             140             145
236
237    CTT GGG TTC CTC AGC ACT GGG GAC GCC AAT CTG CCA GGT AAC TAT GGC    639
238    Leu Gly Phe Leu Ser Thr Gly Asp Ala Asn Leu Pro Gly Asn Tyr Gly
239           150             155             160
240
241    CTT CGG GAT CAG CAC ATG GCC ATT GCT TGG GTG AAG AGG AAT ATC GCG    687
242    Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys Arg Asn Ile Ala
243           165             170             175
244
245    GCC TTC GGG GGG GAC CCC AAC AAC ATC ACG CTC TTC GGG GAG TCT GCT    735
246    Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly Glu Ser Ala
247    180             185             190             195
248
249
250    GGA GGT GCC AGC GTC TCT CTG CAG ACC CTC TCC CCC TAC AAC AAG GGC    783
251    Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr Asn Lys Gly
252                 200             205             210
253
254    CTC ATC CGG CGA GCC ATC AGC CAG AGC GGC GTG GCC CTG AGT CCC TGG    831
255    Leu Ile Arg Arg Ala Ile Ser Gln Ser Gly Val Ala Leu Ser Pro Trp
256                 215             220             225
257
258    GTC ATC CAG AAA AAC CCA CTC TTC TGG GCC AAA AAG GTG GCT GAG AAG    879
```

# SEQUENCE VERIFICATION REPORT
## PATENT APPLICATION   *US/09/418,176*

| Line | Error | Original Text |
|------|-------|---------------|
| 544 | Stop Codon at end of sequence removed - no error | (2) INFORMATION FOR SEQ ID NO:3: |